

**Authors**

Matthew Jukes  
RTI International

**Researchers**

Matthew Jukes  
RTI International

Margaret (Peggy) Dubeck  
RTI International

Elizabeth Turner  
Duke Global Health Institute

Hellen Inyega  
University of Nairobi

Elizabeth Adelman  
Harvard University

Katherine Halliday  
London School of Hygiene and Tropical Medicine

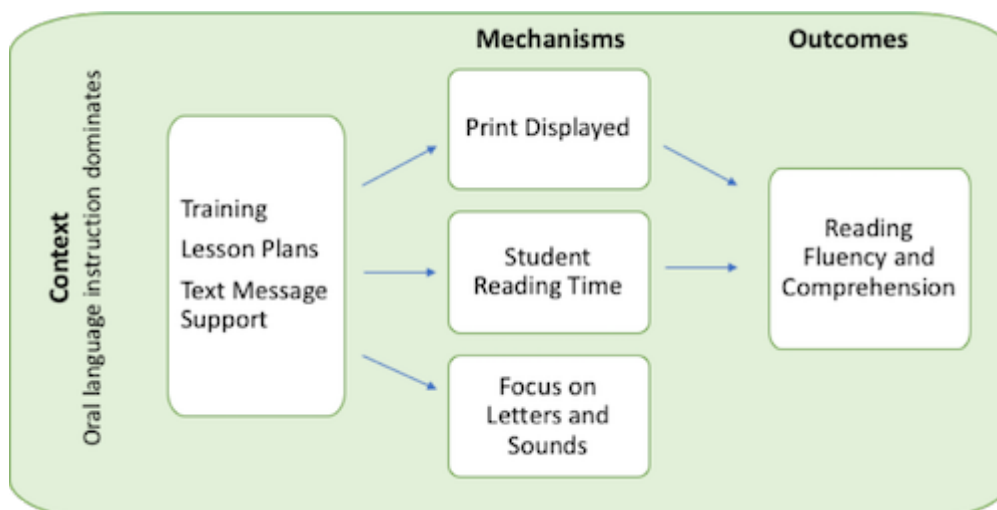
Stephanie Simmons Zuilkowski

Sharon Wolf  
University of Pennsylvania

Simon Brooker  
London School of Hygiene and Tropical Medicine

# Learning More from Impact Evaluations: Contexts, Mechanisms and Theories of Literacy Instruction Interventions

*Editor's note: this is a cross-posting with [the RTI blog](#).*



*Context, mechanisms, and outcomes of the HALI literacy intervention*

If the word “science” in the term “social science” means what I think it means, we social scientists should be in the business of developing models for how we think the world works, testing them with data, and refining them. But the slew of results from impact evaluations in international education frequently feels like a catalogue (if we’re lucky, a meta-analysis) of effect sizes, not a genuine attempt to understand the phenomena we study. To improve our understanding, we need to do a better job of learning from impact evaluations.

The strength of impact evaluations (especially RCTs) is their ability to draw strong causal conclusions about the outcomes that result from a project or program. They can tell you if improvements in learning were *because* of your intervention. The downside is that the causal conclusion only applies to one specific version of program design, implemented in one context. You don’t know as surely whether the project would work elsewhere or in a slightly modified form. It’s important to understand these issues of context and design when formulating evidence-based policy. We need to know if a project will work in different contexts, for example in other countries or other regions of the same country. We need to know if the context of the pilot project involved certain enabling conditions (e.g. strong school management) that need to be in place for the intervention to work. We also need to understand the key aspects of a program design that are critical for its success. This is helpful, for example, if we want to make the program more cost-effective by reducing peripheral components. An understanding of key design elements is particularly crucial when a program is scaled up and embedded in an education system. In this case, some changes to project design are required, for example to adapt to an implementing institution with fewer resources and lower capacity than was available at the pilot.

We need to ensure that the “intervention” will work when it is no longer an intervention but an integral part of the routine activity of an education ministry.

These issues were dissected at the recent inception paper launch of the [Centre for Excellence for Development Impact and Learning \(CEDIL\)](#), a DFID-funded initiative to develop and test innovative methods for evaluation and evidence synthesis. One paper by Calum Davey, James Hargreaves and colleagues<sup>1</sup> provides an excellent framework for thinking about the transport of evidence from one setting to another. They point out that most development interventions aim to achieve “fidelity of form” – implementing a set of components as intended. However, the tension between faithfully replicating a “proven” intervention and adapting it to a new context causes problems. Instead, “fidelity of function” may be a more important focus. This entails ensuring that the intervention has the same effect in the new context. This concept is important for transferring an intervention from one context to another, and particularly when embedding it in a system. However, achieving fidelity of function is only possible when you have an informative theory for how the intervention works. A theory helps us achieve “fidelity of function” – to help us ensure an intervention will have the intended effect – by identifying the *mechanisms* by which the intervention works and characteristics of the *context* that will trigger or enable the mechanism to work. Evaluations are most informative when they report key aspects of the context and mechanisms as well as the outcomes achieved. The context-mechanisms-outcome approach of reporting evaluation

findings has a long history<sup>2</sup> as well as its recent proponents.

## **HALI project and the theory of literacy interventions**

Using the context-mechanisms-outcomes approach, we attempted to add to the theory of literacy instruction interventions in our randomized evaluation of the HALI literacy project. We had previously shown that a program of teacher professional development with text message support had improved children's reading outcomes. We had also identified key aspects of the context for this intervention with pre-baseline classroom observations which told us that oral language was prioritized over interaction with text and breaking down words into letters and sounds. In a recent paper, with Sharon Wolf, Liz Turner, and Peggy Dubeck, we examined the mechanisms by which the literacy intervention acted - what were teachers and students doing differently that led to improved learning? We found that exposure to text - in particular displaying more print in the classroom and giving students more time to read - increased as a result of the intervention and was associated with improved learning. We also found that teachers spent more time focusing on letters and sounds, but this change was not associated with learning outcomes.

The context-mechanisms-outcomes approach allows, I suggest, for a more useful and transportable conclusion from our study. We don't offer a prescriptive conclusion based on fidelity of form, such as "we recommend others do what we did - implement a teacher training intervention with bi-annual workshops, 186 lesson plans and text message support." Instead our conclusion allows others to target fidelity of form: "where oral language instruction is dominant in classrooms (context), a literacy instruction intervention that increases interactions with text (mechanisms) will improve reading skills (outcomes)." In some ways this is a modest contribution to the theory of literacy instruction interventions and experts will not be surprised by our conclusion. But improving literacy instruction is complex and few interventions in low-capacity settings result in teachers implementing everything they learn in training. With this in mind, empirical evidence that helps us focus our efforts is surely welcome.



## **Building theories in international education**

How can the field of international education do a better job of learning from evaluations? One challenge is that most evaluations are driven by accountability. Donors, implementers and other stakeholders (quite rightly) want to know first of all whether the intervention has "worked". But relatively simple extensions of evaluations-for-accountability can help us to learn more. At the most basic level, evaluations can do more to understand whether interventions are implemented with adequate fidelity before trying to understand their impact.

Why devote huge resources to assessing student achievement only to discover teachers were unwilling to implement your intervention in the first place?



*Increasing student exposure to text was a critical factor in the improving student literacy outcomes in the HALI intervention.*

Beyond implementation issues, we can do more to understand the mechanisms by which interventions act and the contexts that enable or trigger them. This can be done through quantitative measures of key processes (as in our paper) or through mixed-method approaches, using qualitative data to understand how interventions worked. We can also set up studies with an explicit aim to test and develop a few elements of a theory, rather than conducting a comprehensive program evaluation. There are other approaches to understanding mechanisms too. Multi-arm trials that systematically vary parameters to understand how things work, case studies and mechanism experiments. The toolbox is large and expanding.

One barrier to the study of mechanisms is that there are many of them. Creating a theory of change for complex social interventions – and the term “complex social intervention” applies to pretty much everything in education – either results in a baffling forest of boxes and arrows or a banal simplification. However, a little pre-baseline research and thinking can help identify key mechanisms to focus on, by asking questions such as “What is the biggest change required in behavior?” and “What changes are likely to make the biggest difference to outcomes?” It helps, too, to have a flexible theory that can be updated as data is collected and experiences of implementers are reported back. It may take a few iterations, but identifying candidate mechanisms should be achievable in any evaluation.

If you needed convincing, I hope I have persuaded you that we need to better understand how things work-- not just if they work-- and ensure that we have the tools to do it. In addition to providing information for accountability, we need to probe further, into the black box of mechanisms. If we care about our evaluations, we need to look them in the eye and tell them they are more than just an effect size. We need to be theory-building, hypothesis-testing, curious scientists. This orientation should ensure that evaluations produce genuine, generalizable, usable knowledge. Based on an understanding of key mechanisms and the contexts that trigger them, program designs could be more effective and better adapted to local conditions. In short, theory-based evaluations improve practice. At least in theory.

[1] Davey, C., Hargreaves, J., Hassan, S., Cartwright, N., Humphreys, M., Masset, E., Prost, A., Gough, D., Oliver, S., & Bonell, C. (2018). Designing evaluations to provide evidence to inform action in new settings. Paper presented at the CEDIL (Centre for Excellence in Development Impact and Learning) Inception Paper Launch, London, UK.

[2] Pawson, R. and Tilley, N. (1997) Realistic Evaluation, London: SAGE

April 03, 2018